

WEIGHTED REGULARITY LEMMA WITH APPLICATIONS

BÉLA CSABA AND ANDRÁS PLUHÁR

ABSTRACT. We prove an extension of the Regularity Lemma with vertex and edge weights, which can be applied for a large class of sparse graphs.

1. INTRODUCTION

Let $G = G(A, B)$ be a bipartite graph. For $X, Y \subset A \cup B$ let $e_G(X, Y)$ denote the number of edges with one endpoint in X and the other in Y . We say that the (A, B) -pair is ε -regular if

$$\left| \frac{e_G(A', B')}{|A'| |B'|} - \frac{e_G(A, B)}{|A| |B|} \right| < \varepsilon$$

for every $A' \subset A$, $|A'| > \varepsilon |A|$ and $B' \subset B$, $|B'| > \varepsilon |B|$.

This definition plays a crucial role in the celebrated Regularity Lemma of Szemerédi, see [8, 9]. Szemerédi invented the lemma for proving his famous theorem on arithmetic progression of dense subsets of the natural numbers. The Regularity Lemma is a very powerful tool when applied to a dense graph. It has found lots of applications in several areas of mathematics and computer science, for applications in graph theory see e.g. [7]. However, it does not tell us anything useful when applied for a sparse graph (i.e., a graph on n vertices having $o(n^2)$ edges).

There has been significant interest to find widely applicable versions for sparse graphs. This turns out to be a very hard task. Kohayakawa [5] proved a sparse regularity lemma, and with Rödl and Łuczak [6] they applied it for finding arithmetic progressions of length 3 in dense subsets of a random set. In their sparse regularity lemma dense graphs are substituted by dense subgraphs of a random (or quasi-random) graph. Naturally, a new definition of ε -regularity was needed, below we formulate a slightly different version from theirs:

Let $F(A, B)$ and $G(A, B)$ be two bipartite graphs such that $F \subset G$. We say that the (A, B) -pair is ε -regular in F relative to G if

$$\left| \frac{e_F(A', B')}{e_G(A', B')} - \frac{e_F(A, B)}{e_G(A, B)} \right| < \varepsilon$$

Key words and phrases. Regularity lemma, sparse graphs.

The authors were partially supported by the grants OTKA T049398 and OTKA K76099. The first author was partially supported by a New Faculty Scholarship Grant of the Office of Sponsored Programs at Western Kentucky University.

for every $A' \subset A, B' \subset B$ and $|A'| > \varepsilon|A|, |B'| > \varepsilon|B|$. It is easy to see that the above is a generalization of ε -regularity, in the original definition the role of G is played by the complete bipartite graph $K_{A,B}$. In this more general definition F can be a rather sparse graph, it only has to be dense *relative to* G , that is, $e(F)/e(G)$ should be a constant.

In this paper we further generalize the notion of quasi-randomness and ε -regularity by introducing *weighted regularity* using vertex and edge weights. This enables us to prove a more general, and perhaps more applicable regularity lemma. Let us remark, that a new notion, *volume regularity* was recently used by Alon et al. [2]. As we will see later, volume regularity is a special case of weighted regularity. Still, these approaches are not fully comparable, since they use different premises for finding the regular partition of a graph. In this paper we discuss the Regularity Lemma for dense subgraphs of *weighted quasi-random* graphs, while Alon et al. find a regular partition for so called (C, η) -bounded graphs.

In the proof we will use the Strong Structure Theorem of Tao [10], that simplifies the proof and allows a shorter exposition. We remark that similar ideas might be used to find a regularity lemma for certain sparse hypergraphs as well.

The structure of the paper is as follows. First we discuss weighted quasi-randomness and weighted ε -regularity in the second section. In the third section we prove the new version of the regularity lemma. Finally, we show some applications in the fourth section.

2. BASIC DEFINITIONS AND TOOLS

Let $\beta > 0$ and $G = (V, E)$ be a graph on n vertices. Set $\delta_G = e(G)/\binom{n}{2}$, this is the *density* of G . We define the density of the A, B pair of subsets of $V(G)$ by $\delta_G(A, B) = e_G(A, B)/(|A||B|)$. We say that G is β -quasi-random if it has the following property: If $A, B \subset V(G)$ such that $A \cap B = \emptyset$ and $|A|, |B| > \beta n$ then

$$|\delta_G - \delta_G(A, B)| < \beta \delta_G.$$

That is, the edges of G are distributed “randomly.” In order to formulate our regularity lemma we have to define quasi-randomness in a more general way, that admits weights on vertices and edges.

For a function $w : S \rightarrow \mathbb{R}^+$ and $A \subset S$, $w(A)$ is defined by the usual way, that is, $w(A) = \sum_{x \in A} w(x)$. We shall also use the indicator function of the edge set of a graph H . $\mathbf{1}_H : \binom{V(H)}{2} \rightarrow \{0, 1\}$ and $\mathbf{1}_H(x, y) = 1$ iff $xy \in E(H)$.

We define the weighted quasi-randomness of a graph $G = (V, E)$ with given weight-functions $\mu : V \rightarrow \mathbb{R}^+$ and $\rho : \binom{V}{2} \rightarrow \mathbb{R}^+$. For $A, B \subset V$ let

$$\rho_G(A, B) := \sum_{u \in A, v \in B} \mathbf{1}_G(u, v) \rho(u, v).$$

In particular, $\rho_G(u, v) = \mathbf{1}_G(u, v) \rho(u, v)$ for $u, v \in V$.

Definition 1. A graph $G = (V, E)$ is weighted β -quasi-random with weight-functions μ and ρ if for any $A, B \subset V(G)$ such $A \cap B = \emptyset$ and $\mu(A) \geq \beta\mu(V)$, $\mu(B) \geq \beta\mu(V)$ we have

$$\left| \frac{\rho_G(A, B)}{\mu(A)\mu(B)} - \frac{\rho_G(V, V)}{\mu(V)\mu(V)} \right| < \beta.$$

Observe that choosing $\mu \equiv 1$ and $\rho \equiv 1/\delta_G$ gives back the first definition of quasi-randomness. There is another, weaker notion of quasi-randomness, which we will also use.

Definition 2. Let $D > 1$ be an absolute constant. A graph $G = (V, E)$ is weighted (D, β) -quasi-random with weight-functions μ and ρ if for any $A, B \subset V(G)$ such $A \cap B = \emptyset$ and $\mu(A) \geq \beta\mu(V)$, $\mu(B) \geq \beta\mu(V)$ we have

$$\frac{1}{D} \frac{\rho_G(V, V)}{\mu(V)\mu(V)} \leq \frac{\rho_G(A, B)}{\mu(A)\mu(B)} \leq D \frac{\rho_G(V, V)}{\mu(V)\mu(V)}.$$

Clearly, if a graph is β -quasi-random, then it is (D, β) -quasi-random unless D is very close to one. Now we need to describe the weighted version of relative regularity.

Definition 3. Let G and F be graphs such that $F \subset G$ and μ, ρ weight functions defined as above. For $A, B \subset V(G)$ and $A \cap B = \emptyset$ the pair (A, B) in F is $\mu - \rho$ -weighted ϵ -regular relative to G , or briefly weighted ϵ -regular, if

$$\left| \frac{\rho_F(A', B')}{\mu(A')\mu(B')} - \frac{\rho_F(A, B)}{\mu(A)\mu(B)} \right| < \epsilon$$

for every $A' \subset A$ and $B' \subset B$ provided that $\mu(A') \geq \epsilon\mu(A)$, $\mu(B') \geq \epsilon\mu(B)$. Here

$$\rho_F(A, B) = \sum_{u \in A, v \in B} \mathbf{1}_F(u, v) \rho(u, v).$$

Remark. Note that weighted ϵ -regularity is nothing but the well-known ϵ -regularity when $G = K_{A, B}$ and $\mu \equiv 1$ and ρ is chosen to be identically the reciprocal of the density of G as before.

Next we define weighted regular partitions.

Definition 4. Let $G = (V, E)$ and $F \subset G$ be graphs, and μ and ρ weight functions. F has a weighted ϵ -regular partition relative to G if its vertex set V can be partitioned into $\ell + 1$ clusters W_0, W_1, \dots, W_ℓ such that

- $\mu(W_0) \leq \epsilon\mu(V)$,
- $|\mu(W_i) - \mu(W_j)| \leq \max_{x \in V} \{\mu(x)\}$ for every $1 \leq i, j \leq \ell$,
- all but at most $\epsilon\ell^2$ of the pairs (W_i, W_j) for $1 \leq i < j \leq \ell$ are weighted ϵ -regular in F relative to G .

In order to show our main result we will use the Strong Structure Theorem of Tao, that allows a short exposition. In fact we will closely follow his proof for the Regularity Lemma as discussed in [10]. First we have to introduce some definitions.

Let H be a real finite-dimensional Hilbert space, and let S be a set of basic functions or basic structured vectors of H of norm at most 1. The function $g \in H$ is (M, K) -structured with the positive integers M, K if one has a decomposition

$$g = \sum_{1 \leq i \leq M} c_i s_i$$

with $s_i \in S$ and $c_i \in [-K, K]$ for $1 \leq i \leq M$. We say that g is β -pseudorandom for some $\beta > 0$ if $|\langle g, s \rangle| \leq \beta$ for all $s \in S$. Then we have the following

Theorem 1 (Strong Structure Theorem - T. Tao). *Let H and S be as above, let $\varepsilon > 0$, and let $J : \mathbf{Z}^+ \rightarrow \mathbb{R}^+$ be an arbitrary function. Let $f \in H$ be such that $\|f\|_H \leq 1$. Then we can find an integer $M = M_{J, \varepsilon}$ and a decomposition $f = f_{str} + f_{psd} + f_{err}$ where (i) f_{str} is (M, M) -structured, (ii) f_{psd} is $1/J(M)$ -pseudorandom, and (iii) $\|f_{err}\|_H \leq \varepsilon$.*

3. WEIGHTED REGULARITY LEMMA RELATIVE TO A QUASI-RANDOM GRAPH G

First we define the Hilbert space H , and S . We generalize Example 2.3 of [10] to weighted graphs. Let $G = (V, E)$ be a β -quasi-random graph on n vertices with weight functions μ and ρ . Let H be the $\binom{n}{2}$ -dimensional space of functions $g : \binom{V}{2} \rightarrow \mathbb{R}$, endowed with the inner product

$$\langle g, h \rangle = \frac{1}{\binom{n}{2}} \sum_{(u,v) \in \binom{V}{2}} g(u,v) h(u,v) \rho_G(u,v).$$

It is useful to normalize the vertex and edge weight functions, we assume that $\mu(V) = n$ and $\langle 1, 1 \rangle = 1$. We also assume, that $\mu(v) = o(|V|)$ for every $v \in V$. Observe, that if $F \subset G$ then $\|\mathbf{1}_F\| \leq 1$. We let S to be the collection of 0,1-valued functions $\gamma_{A,B}$ for $A, B \subset V(G)$, $A \cap B = \emptyset$, where $\gamma_{A,B}(u,v) = 1$ if and only if $u \in A$ and $v \in B$. We have the following

Theorem 2 (Weighted Regularity Lemma). *Let $D > 1$ and $\beta, \varepsilon \in (0, 1)$, such that $0 < \beta \ll \varepsilon \ll 1/D$ and let $L \geq 1$. If $G = (V, E)$ is a weighted (D, β) -quasi-random graph on n vertices with n sufficiently large depending on ε and L and $F \subset G$, then F admits a weighted ε -regular partition relative to G into the partition sets W_0, W_1, \dots, W_ℓ such that $L \leq \ell \leq C_{\varepsilon, L}$ for some constant $C_{\varepsilon, L}$.*

Proof: Let us apply Theorem 1 to the function $\mathbf{1}_F$ with parameters η and function J to be chosen later. We get the decomposition

$$\mathbf{1}_F = f_{str} + f_{psd} + f_{err},$$

where f_{str} is (M, M) -structured, f_{psd} is $1/J(M)$ -pseudorandom, and $\|f_{err}\| \leq \eta$ with $M = M_{J, \eta} = M_{J, \varepsilon}$.

The function f_{str} is the combination of at most M basic functions:

$$f_{str} = \sum_{1 \leq k \leq M} \alpha_k \gamma_{\mathcal{A}_k, \mathcal{B}_k}$$

where $\mathcal{A}_k, \mathcal{B}_k$ are subsets of V and $\gamma_{\mathcal{A}_k, \mathcal{B}_k}$ agrees with the indicator function of the edges of G in between \mathcal{A}_k and \mathcal{B}_k . Any $(\mathcal{A}_k, \mathcal{B}_k)$ pair partitions V into at most 4 subsets. Overall we get a partitioning of V into at most 4^M subsets, we will refer to them as *atoms*. Divide every atom into subsets of total vertex weight $\frac{\varepsilon n}{L+4^M}$, except possibly one smaller subset. The small subsets will be put into W_0 , the others give W_1, W_2, \dots, W_ℓ , with $\ell = \frac{L+4^M}{\varepsilon}$. We refer to the sets W_i for $i = 1, \dots, \ell$ as *clusters*. If n is sufficiently large then this partitioning is non-trivial. From the construction it follows that each W_i is entirely contained within an atom. It is also clear that $\mu(W_0) \leq \varepsilon n$ and $\mu(W_i) \approx m = \Theta(\frac{n}{\ell})$ for every $1 \leq i \leq \ell$.

We have that

$$\|f_{err}\|^2 = \frac{1}{\binom{n}{2}} \sum_{(u,v) \in \binom{V}{2}} |f_{err}(u,v)|^2 \rho_G(u,v) \leq \eta^2.$$

From this and the normalization of ρ it follows that

$$\frac{1}{\binom{\ell}{2}} \sum_{1 \leq i < j \leq \ell} \frac{1}{\rho_G(W_i, W_j)} \sum_{u \in W_i, v \in W_j} |f_{err}(u,v)|^2 \rho_G(u,v) = O(\eta^2).$$

Clearly,

$$\frac{1}{\rho_G(W_i, W_j)} \sum_{u \in W_i, v \in W_j} |f_{err}(u,v)|^2 \rho_G(u,v) = O(\eta)$$

for all but at most $O(\eta \ell^2)$ pairs (i, j) . If the above is satisfied for a pair (i, j) then we call it a *good pair*. We will apply the Cauchy-Schwarz inequality. For that let $a(u,v) = |f_{err}(u,v)| \sqrt{\rho_G(u,v)}$ and $b(u,v) = \sqrt{\rho_G(u,v)}$, then

$$\frac{\sum_{u \in W_i, v \in W_j} a(u,v)b(u,v)}{\sqrt{\sum_{u \in W_i, v \in W_j} b^2(u,v)}} \leq \sqrt{\sum_{u \in W_i, v \in W_j} a^2(u,v)}.$$

Since

$$\sqrt{\sum_{u \in W_i, v \in W_j} a^2(u,v)} = O(\sqrt{\eta}) \sqrt{\rho_G(W_i, W_j)},$$

we get that

$$\frac{1}{\rho_G(W_i, W_j)} \sum_{u \in W_i, v \in W_j} |f_{err}(u,v)| \rho_G(u,v) = O(\sqrt{\eta})$$

if (i, j) is a good pair.

Assume that (i, j) is a good pair. From the pseudorandomness of f_{psd} we have that

$$|\langle f_{psd}, \gamma_{A,B} \rangle| = \frac{1}{\binom{n}{2}} \left| \sum_{u \in A, v \in B} f_{psd}(u, v) \rho_G(u, v) \right| \leq \frac{1}{J(M)}$$

for every $A \subset W_i$ and $B \subset W_j$.

We will show that every good pair is weighted ε -regular in F relative to G . Let (i, j) be a good pair, and assume that $A \subset W_i$, $\mu(A) > \varepsilon \mu(W_i)$ and $B \subset W_j$, $\mu(B) > \varepsilon \mu(W_j)$. If (W_i, W_j) is weighted ε -regular, then

$$\left| \frac{\rho_F(A, B)}{\mu(A)\mu(B)} - \frac{\rho_F(W_i, W_j)}{\mu(W_i)\mu(W_j)} \right| < \varepsilon.$$

Recall that

$$\rho_F(A, B) = \sum_{u \in A, v \in B} \mathbf{1}_F(u, v) \rho(u, v) = \sum_{u \in A, v \in B} \mathbf{1}_F(u, v) \rho_G(u, v),$$

since $F \subset G$.

Substituting $f_{str} + f_{psd} + f_{err}$ for $\mathbf{1}_F$ it is sufficient to verify the following inequalities.

$$(1) \left| \frac{\sum_{u \in A, v \in B} f_{str}(u, v) \rho_G(u, v)}{\mu(A)\mu(B)} - \frac{\sum_{u \in W_i, v \in W_j} f_{str}(u, v) \rho_G(u, v)}{\mu(W_i)\mu(W_j)} \right| < \varepsilon/3,$$

$$(2) \left| \frac{\sum_{u \in A, v \in B} f_{psd}(u, v) \rho_G(u, v)}{\mu(A)\mu(B)} - \frac{\sum_{u \in W_i, v \in W_j} f_{psd}(u, v) \rho_G(u, v)}{\mu(W_i)\mu(W_j)} \right| < \varepsilon/3$$

and

$$(3) \left| \frac{\sum_{u \in A, v \in B} f_{err}(u, v) \rho_G(u, v)}{\mu(A)\mu(B)} - \frac{\sum_{u \in W_i, v \in W_j} f_{err}(u, v) \rho_G(u, v)}{\mu(W_i)\mu(W_j)} \right| < \varepsilon/3.$$

For proving (1) recall that f_{str} is constant on $W_i \times W_j$ and (M, M) -structured. Since the $\gamma_{X,Y}$ basic functions are 0, 1-valued, we get, that $|f_{str}| \leq M^2$. Moreover, G is (D, β) -quasi-random, where $0 < \beta \ll \varepsilon$. Therefore, (1) $\leq DM^2\beta < \varepsilon/3$, since $\beta \ll \varepsilon$.

The proof of (2) goes as follows. The first term is

$$\left| \frac{\sum_{u \in A, v \in B} f_{psd}(u, v) \rho_G(u, v)}{\mu(A)\mu(B)} \right| = \binom{n}{2} |\langle f_{psd}, \gamma_{A,B} \rangle| \leq \frac{\binom{n}{2}}{J(M)\mu(A)\mu(B)}$$

the second is

$$\left| \frac{\sum_{u \in W_i, v \in W_j} f_{psd}(u, v) \rho_G(u, v)}{\mu(W_i)\mu(W_j)} \right| = \binom{n}{2} |\langle f_{psd}, \gamma_{W_i, W_j} \rangle| \leq \frac{\binom{n}{2}}{J(M)\mu(W_i)\mu(W_j)}.$$

Noting that $\mu(W_k) = \Theta(n/\ell)$ for $k \geq 1$ we get that the sum of the above terms is at most

$$\frac{\ell^2}{2J(M)} \left(1 + \frac{1}{\varepsilon^2}\right) < \frac{\varepsilon}{3},$$

if $J(M) \gg \frac{\ell^2}{\varepsilon^3}$.

For (3) first notice that it is upper bounded by

$$O(\sqrt{\eta}) \left(\frac{\rho_G(W_i, W_j)}{\mu(W_i)\mu(W_j)} + \frac{\rho_G(W_i, W_j)}{\mu(A)\mu(B)} \right) \leq O(\sqrt{\eta}) \frac{\rho_G(W_i, W_j)}{\varepsilon^2 \mu(W_i)\mu(W_j)}.$$

We also have that

$$\frac{\rho_G(W_i, W_j)}{\mu(W_i)\mu(W_j)} = O(1)$$

by the normalization of μ and ρ and from the fact that G is quasi-random. From this it is easy to see that if $\eta \ll \varepsilon^6$ then (3) is at most $\varepsilon/3$. This finishes the proof of the theorem. \square

4. APPLICATIONS

In this section we consider a few applications of our main result. First we prove that a random graph with widely differing edge probabilities is quasi-random, if none of the edge probabilities are too small. In this case the vertex weights will all be one, but edges will have different weights. Then we show examples where vertices have different weights. We will also consider the relation of weighted regularity and volume regularity.

4.1. Quasi-randomness in the $G(n, p_{ij})$ model. In this section we will prove that random graphs of the $G(n, p_{ij})$ model are quasi-random in the strong sense with high probability. A special case of this model is the well-known $G(n, p)$ model for random graphs. A Regularity Lemma for this case was first applied by Kohayakawa, Łuczak and Rödl in [6]. They studied $G(n, p)$ for $p = c/\sqrt{n}$ in order to find arithmetic progressions of length three in dense subsets of random subsets of $[N]$.

The $G(n, p_{ij})$ model was first considered by Bollobás [4]. Recently it was also studied by Chung et al. [1].

In this model one takes n vertices, and draws an edge between the vertices x_i and x_j with probability p_{ij} , randomly and independently of each other. Note that if $p_{ij} \equiv p$, then we get back the well-known $G(n, p)$ model. It is a straightforward application of the Chernoff bound that a random graph $G \in G(n, p)$ is quasi-random with high probability. However, the case of $G(n, p_{ij})$ is somewhat harder.

Lemma 3. *Let $\beta > 0$. There exists a $K = K(\beta)$ such that if $G \in G(n, p_{ij})$ and $p_{ij} \geq K/n$ for every i and j , then G is weighted β -quasi-random with probability at least $1 - 2^{-n}$ if n is sufficiently large.*

Proof. First of all let $\mu \equiv 1$, and let $\rho(i, j) = 1/p_{i,j}$. Set $K = 4800/\beta^6$. Let $p_0 = K/n$, and let $p_k = e^k p_0$ for $1 \leq k \leq \log n$. Let A and B be a pair of disjoint sets, both of size at least βn . We partition the pairs (u, v) , where $u \in A$ and $v \in B$, into $O(\log n)$ disjoint sets H_1, H_2, \dots, H_l : if $p_k \leq p_{uv} < p_{k+1}$ then (u, v) will belong to H_k . Let $a_k = \frac{\beta^3}{10} \sqrt{e^k} K n$. We will denote $|H_k|$ by m_k .

We will prove that the following inequality holds with probability at least $1 - 2^{-3n}$:

$$\left| \sum_{u \in A, v \in B} \frac{X_{uv}}{p_{uv}|A||B|} - 1 \right| < \beta/2,$$

where X_{uv} is a random variable which is 1 if $uv \in E(G)$, otherwise it is 0. This implies the quasi-randomness of G since there are less than 2^{2n} pairs of disjoint subsets of $V(G)$. Observe that

$$\sum_{u \in A, v \in B} \frac{\mathbb{E}X_{uv}}{p_{uv}|A||B|} = 1.$$

Applying the large deviation inequalities A.1.11 and A.1.13 from [3], we are able to bound the number of edges in between A and B for the edges of H_k in case m_k is sufficiently large as follows. According to A.1.11 we have that

$$\Pr \left(\sum_{(u,v) \in H_k} (X_{uv} - \mathbb{E}X_{uv}) > a_k \right) < e^{-\frac{a_k^2}{2q_k m_k} + \frac{a_k^3}{2q_k^2 m_k^2}},$$

where

$$p_k \leq q_k = \sum_{(u,v) \in H_k} \frac{p_{uv}}{m_k} < p_{k+1}.$$

We estimate the exponent in case $m_k = n^2$:

$$-\frac{a_k^2}{2q_k m_k} + \frac{a_k^3}{2q_k^2 m_k^2} \leq -\frac{\beta^6}{200} \left(\frac{\sqrt{e^2}}{e} \right)^k \frac{K n^3}{e m_k} + \frac{\beta^9}{2000} \left(\frac{\sqrt{e^3}}{e^2} \right)^k \frac{e K n^5}{m_k^2} < -3n,$$

where we used the definition of K . For m_k being much less than n^2 direct substitution gives a useless bound. For this case we have the useful inequality

$$\frac{1}{2} \Pr \left(\sum_{i=1}^{m_k} Y_i > a_k \right) \leq \Pr \left(\sum_{i=1}^{n^2} Y_i > \frac{a_k}{2} \right),$$

where $\Pr(Y_i = 1 - q_k) = q_k$ and $\Pr(Y_i = -q_k) = 1 - q_k$. This implies that the exponent is at most $-3n$ even in case $m_k < n^2$.

Indeed, let A, B and C be the events that $\sum_{i=1}^{m_k} Y_i > a_k$, $\sum_{i=1}^{n^2} Y_i > a_k/2$ and $\sum_{i=m_k+1}^{n^2} Y_i < -a_k/2$, respectively. Clearly A and C are independent, and $A \cap \overline{C} \subset B$.

So we have $\Pr(B) \geq \Pr(A \cap \overline{C}) = \Pr(A)\Pr(\overline{C})$, that is $\Pr(A) \leq \Pr(B)/\Pr(\overline{C}) < \Pr(B)/2$, since by A.1.13

$$\Pr\left(\sum_{i=m_k+1}^{n^2} Y_i < -\frac{a_k}{2}\right) < e^{-\frac{a_k^2}{8q_k(n^2-m_k)}} < \frac{1}{2}.$$

With this we have proved that the sum of the weights of the edges of H_k will not be much larger than their expectation with high probability.

Now we estimate the probability that the sum of the weights is much less than their expectation. Let us use A.1.13 again directly to the sums over H_k 's:

$$\Pr\left(\sum_{(u,v) \in H_k} (X_{uv} - \mathbb{E}X_{uv}) < -a_k\right) < e^{-\frac{a_k^2}{2q_k m_k}}.$$

The exponent in the inequality can be estimated very similarly as before:

$$-\frac{a_k^2}{2q_k m_k} \leq -\frac{\beta^6}{200} \left(\frac{\sqrt{e^2}}{e}\right)^k \frac{Kn^3}{em_k} < -3n,$$

moreover, this bound applies for an arbitrary m_k .

Putting these together we have that

$$\Pr\left(\left|\sum_{(u,v) \in H_k} (X_{uv} - \mathbb{E}X_{uv})\right| > a_k\right) < 2^{-3n}.$$

This implies that

$$\left|\sum_{(u,v) \in H_k} \frac{X_{uv} - \mathbb{E}X_{uv}}{p_{uv}|A||B|}\right| \leq \left|\sum_{(u,v) \in H_k} \frac{X_{uv} - \mathbb{E}X_{uv}}{p_{k-1}|A||B|}\right| \leq \frac{a_k}{p_{k-1}|A||B|} \leq \frac{\beta}{10} \left(\frac{1}{\sqrt{e}}\right)^k,$$

where the last two inequalities hold with probability at least $1 - 2^{-3n}$ for a given pair of sets A and B if n is sufficiently large. Since

$$\left|\sum_{u \in A, v \in B} \frac{X_{uv}}{p_{uv}|A||B|}\right| = \left|\sum_{k=1}^{\log n} \sum_{(u,v) \in H_k} \frac{X_{uv}}{p_{uv}|A||B|}\right| \leq \sum_{k=1}^{\log n} \left|\sum_{(u,v) \in H_k} \frac{X_{uv}}{p_{k-1}|A||B|}\right|$$

and

$$\sum_{k=1}^{\log n} \frac{1}{10} \left(\frac{1}{\sqrt{e}}\right)^k \leq \frac{1}{2},$$

the claimed bound follows with high probability. \square

Remark. It is very similar to prove that with high probability $|\sum_{i,j} \rho_G(i,j) - \binom{n}{2}| = o(n)$, we omit the details. From this it follows that rescaling the above edge weights by a factor of $(1+o(1))$ and letting $\mu \equiv 1$ provides us β -quasi-random weights for most graphs from $G(n, p_{ij})$ such that $\mu(V) = n$ and $\rho_G(V, V) = 2\binom{n}{2}$. That is, with high probability we can apply the Regularity Lemma for any $F \subset G$, where $G \in G(n, p_{ij})$.

4.2. Other examples for defining vertex and edge weights. When defining the notion of weighted quasi-randomness and weighted regularity, we mentioned, that choosing $\mu \equiv 1$ and $\rho \equiv 1/\delta_G$ gives back the old definitions of quasi-randomness and regularity. In the previous section we saw an example when we needed different edge weights, but μ was identically one.

Let us consider a simple example in which μ has to take more than one value. Let G be a star on n vertices, that is, the vertex v_1 is adjacent to the vertices v_2, \dots, v_n , and v_i has degree 1 for $i \geq 2$. We let $\mu(v_1) = 1/2$ and $\mu(v_i) = 1/(2(n-1))$ for $i \geq 2$, and choose $\rho_G \equiv n/2$. With these choices G is easily seen to be quasi-random, moreover, it is weighted regular.

A more sophisticated example relates weighted regularity with volume regularity, the latter introduced by Alon et al [2]. Given a graph G , the volume of a set $S \subset V(G)$ is defined as $\text{vol}(S) = \sum_{v \in S} \deg(v)$. For disjoint sets $A, B \subset V(G)$, the (A, B) pair is ε -volume regular if $\forall X \subset A, Y \subset B$ satisfying $\text{vol}(X) \geq \varepsilon \cdot \text{vol}(A)$ and $\text{vol}(Y) \geq \varepsilon \cdot \text{vol}(B)$ we have

$$\left| e(X, Y) - \frac{e(A, B)}{\text{vol}(A)\text{vol}(B)} \text{vol}(X)\text{vol}(Y) \right| < \varepsilon \cdot \text{vol}(A)\text{vol}(B)/\text{vol}(V).$$

In order to show that volume regularity is a special case of weighted regularity, we may choose edge and vertex weights as follows: for all $A, B \subset V$ let

$$\rho(A, B) = e(A, B) \frac{n^2}{\text{vol}(V)}$$

and

$$\mu(A) = \frac{n \cdot \text{vol}(A)}{\text{vol}(V)}.$$

It is easy to see that if the (A, B) pair is $\mu - \rho$ -weighted ε -regular with the above choice for μ and ρ , then it is also ε -volume regular.

In certain cases it can be more useful to apply the notion of weighted regularity than volume regularity. As an example, consider a bipartite graph $G(A, B)$, where $A = A_1 \cup A_2$, $B = B_1 \cup B_2$ and $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$. We assume that $|A_1| = |A_2| = |B_1| = |B_2| = n/4$. Further assume that $G(A_1, B)$ and $G(A, B_1)$ are complete bipartite graphs, and $G(A_2, B_2)$ is a random bipartite graph with edge probability $1/\sqrt{n}$.

Letting $X = A_2$ and $Y = B_2$, it is easy to see that $G(A, B)$ is not ε -volume regular, if ε is small, no matter how large n is. On the other hand, the result of the previous section implies that G is weighted regular, even with a very small ε .

Acknowledgment The authors thank Péter Hajnal and Endre Szemerédi for the helpful discussions.

REFERENCES

- [1] W. Aiello, F. Chung and L. Lu, A random graph model for massive graphs, *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, (2000) 171-180.
- [2] N. Alon, A. Coja-Oghlan, H. Hàn, M. Kang, V. Rödl and M. Schacht, Quasi-Randomness and Algorithmic Regularity for Graphs with General Degree Distributions, *Proc. ICALP 2007*, 789–800.
- [3] N. Alon and J. Spencer, *The Probabilistic Method*. Wiley-Interscience, New York, 2000.
- [4] B. Bollobás, *Random Graphs*. Second edition. Cambridge Studies in Advanced Mathematics, **73**. Cambridge University Press, Cambridge, 2001.
- [5] Y. Kohayakawa, Szemerédi’s Regularity Lemma for sparse graphs, in *Foundations of computational mathematics* (1997) 216–230.
- [6] Y. Kohayakawa, T. Łuczak and V. Rödl, Arithmetic progressions of length three in subsets of a random set, *Acta Arith.* **75** (1996), no. 2, 133–163
- [7] J. Komlós, M. Simonovits, Szemerédi’s Regularity Lemma and its Applications in Graph Theory, *Combinatorics, Paul Erdős is eighty, Vol. 2* (Keszthely, 1993), 295–352.
- [8] E. Szemerédi, On sets of integers containing no k elements in arithmetic progressions, *Acta Arithmetica* **27** (1975), 299–345.
- [9] E. Szemerédi, Regular Partitions of Graphs, *Colloques Internationaux C.N.R.S N^o 260 - Problèmes Combinatoires et Théorie des Graphes*, Orsay (1976), 399-401.
- [10] T. Tao, Structure and randomness in combinatorics, *FOCS’07*

DEPARTMENT OF MATHEMATICS, WESTERN KENTUCKY UNIVERSITY
E-mail address: `bela.csaba@wku.edu`

UNIVERSITY OF SZEGED, DEPARTMENT OF COMPUTER SCIENCE, ÁRPÁD TÉR 2., SZEGED,
 H-6720
E-mail address: `pluhar@inf.u-szeged.hu`